THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

April 8, 1999

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: ***60/086,074***
FILING DATE: ***May 20, 1998***
PCT APPLICATION NUMBER: ***PCT/US99/06797***

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

M. Montgomery

WANDA MONTGOMERY
Certifying Officer

A/PROV

Modified PTO/SB/16 (6-95)
Approved for use through 04/11/98. OMB 0651-0037
Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53 (b)(2).

Docket Number	P-11586A	Type a plus sign (+) inside this box -->	+

05/20/98 JC535 U.S. PTO

INVENTOR(s)/APPLICANT(s)

LAST NAME	FIRST NAME	MIDDLE NAME	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
Dou Song	Shenshen Ho	Yeong	Carmel, Indiana Carmel, Indiana

TITLE OF THE INVENTION (280 characters max)

TUMOR NECROSIS FACTOR RECEPTOR FAMILY PROTEIN AND RELATED NUCLEIC ACID

CORRESPONDENCE ADDRESS

Eli Lilly and Company
Patent Division/D.C. 1104
Lilly Corporate Center
Indianapolis, Indiana 46285

STATE	IN	ZIP CODE	46285	COUNTRY	USA

ENCLOSED APPLICATION PARTS (*check all that apply*)

X	Specification	Number of pages	57		Small Entity Statement		
	Drawing(s)	Number of Sheets			Other (Specify)		

METHOD OF PAYMENT (*check one*)

	A check or money order is enclosed to cover the Provisional filing fees	PROVISIONAL FILING FEE AMOUNT ($)	$150.00
X	The Assistant Commissioner is hereby authorized to charge filing fees and credit Deposit Account Number: 05-0840		

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

X No.

☐ Yes, the name of the U.S. Government agency and the Government contract number are:

Respectfully submitted,
SIGNATURE Paul R Cantrell Date 5 / 20 / 97

TYPED or PRINTED NAME PAUL R. CANTRELL REGISTRATION NO. (*if appropriate*) 36,470

☐ Additional inventors are being named on separately numbered sheets attached hereto

PROVISIONAL APPLICATION FOR PATENT FILING ONLY

"Express Mail" mailing label number EM564070039US Date of Deposit 20 MAY 1998
I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Printed Name Signature

Box - Sept.

P-11586A

TUMOR NECROSIS FACTOR RECEPTOR FAMILY PROTEIN AND RELATED NUCLEIC ACID COMPOUNDS

BACKGROUND OF THE INVENTION

This invention relates to a novel gene and its cognate protein, the protein putatively being a member of the tumor necrosis factor receptor (TNFR) superfamily. Also contemplated are methods for identifying compounds that bind said receptor, and methods for inhibiting osteoclast differentiation and bone resorption.

The TNFR superfamily is a group of type I proteins (generally transmembrane) that share a conserved cysteine-rich motif, which is repeated three to six times in the extracellular domain (Smith, et al., 1994, Cell 76:953-62). Collectively, these repeat units form the ligand binding domains of these receptors (Chen et al., 1995, Chemistry 270:2874-78). The TNFR's are variably expressed in a variety of cell types, including B cells, T cells, dendritic cells, and macrophages.

The ligands for these receptors are a structurally related group of proteins in the tumor necrosis factor (TNF) family. These ligands produce a variety of biological responses in TNFR-bearing cells, including proliferation,

differentiation, immune regulation, inflammatory response, cytotoxicity, and apoptosis, binding to distinct but closely related receptors TNFR-1 and TNFR-2.

Systemic delivery of TNF induces toxic shock and widespread tissue necrosis. Because of this, TNF may be responsible for the severe morbidity and mortality associated with a variety of infectious diseases, including sepsis. Mutations in FasL, the ligand for the TNFR-related receptor FAS/APO (Suda et al., 1993, Cell 75:1169-78, are associated with autoimmunity (Fisher et al., 1995, Cell 81:935-46), while overproduction of FasL may be implicated in drug-induced hepatitis.

Soluble TNFR-1 receptors, and antibodies that bind TNF, have been tested for their ability to neutralize systemic TNFα (Leotsher et al., 1991, Cancer Cells 3(6):221-6). A naturally occurring form of a secreted TNFR-1 mRNA was recently cloned, and its product tested for its ability to neutralize TNF activity *in vitro* and *in vivo* (Kohno et al., 1990, Proc. Nat. Acad, Sci. 87:8331-5).

TNF has also recently been implicated in the pathogenesis of bone loss induced by estrogen deficiency, presumably mediated by binding to certain members of the TNFR superfamily. Expression of a soluble TNFR-1/FcIgG3 fusion protein in transgenic ovariectomized mice was demonstrated to protect against the loss in bone mass and strength experienced by control animals (Ammann et al., 1997, J. Clin. Invest., 99:1699-1703). Moreover, two novel naturally-occurring secreted members of the TNFR superfamily were recently reported as having a role in regulating bone resorption (Simonet et al., 1997, Cell 89:309-19 (termed

osteoprotegerin (OPG)); Tsuda et al., 1997, Biochem. and Biophys. Res. Comm. 234:137-42 (termed osteoclastogenesis inhibitory factor (OCIF))), functioning essentially as inhibitors of differentiation of bone-resorbing osteoclasts.

An object of the present invention is to identify new members of the TNFR superfamily. It is anticipated that new TNFR's may be transmembrane proteins or soluble forms thereof comprising extracellular domains. Indeed, the present invention relates to new nucleic acids and polypeptides encoded thereby that are closely related to TNFR-2, which are implicated in regulation of bone metabolism.

SUMMARY OF THE INVENTION

The present invention provides isolated nucleic acid compounds and novel proteins functionally related to the tumor necrosis factor receptor (TNFR) superfamily. The nucleotide sequences and proteins described herein are referred to as "TNFRsol." TNFRsol protein does not include any transmembrane domains and is, therefore, soluble.

Having the TNFRsol gene enables the production of recombinant TNFRsol protein, the isolation of orthologous genes from other organisms, and/or paralogous genes from the same organism, chromosome mapping studies, and the implementation of large scale screens to identify compounds and potential pharmaceutical agents that bind or regulate expression of said protein and modulate biological activity thereof.

In one embodiment, the present invention relates to an isolated nucleic acid compound encoding TNFRsol

protein, or fragment thereof. A preferred nucleic acid compound comprises the nucleotide sequence identified as SEQ ID NO:1. Other preferred nucleic acid compounds comprise nucleotides 88-900 of SEQ ID NO:1 or nucleotides 102-536 of SEQ ID NO:1.

In another embodiment, the present invention relates to a nucleic acid that is at least 75% identical, and preferably at least 95% identical, to a nucleic acid that encodes SEQ ID NO:2, or fragment thereof.

In another embodiment, the present invention relates to a nucleic acid that hybridizes to the polynucleotide of SEQ ID NO:1, or fragments thereof, under high stringency conditions and encodes a protein that is capable of inhibiting bone resorption.

In another embodiment, the present invention relates to a nucleic acid that hybridizes to the polynucleotide of SEQ ID NO:1, or fragments thereof, under low stringency conditions and encodes a protein that is capable of inhibiting bone resorption.

In another embodiment the present invention relates to an isolated protein molecule, or functional fragment thereof, wherein said protein molecule comprises the sequence identified as SEQ ID NO:2. Examples of functional fragments of preference include amino acid residues 30-300 of SEQ ID NO:2 or residues 34-195 of SEQ ID NO:2.

In yet another embodiment, the present invention relates to a recombinant DNA vector that incorporates SEQ ID NO:1, or fragments thereof, in operable linkage to gene

expression sequences, enabling the gene to be transcribed and translated in a host cell.

In still a further embodiment, the present invention relates to a method for identifying compounds that bind a protein identified herein as SEQ ID NO:2, comprising the steps of admixing a substantially purified preparation of a protein comprising SEQ ID NO:2 with a test compound, and monitoring by any suitable means a binding interaction between said protein and said compound. This method may be employed with peptide fragments of SEQ ID NO:2

This invention also provides a method of determining whether a nucleic acid sequence of the present invention, or fragment thereof, is present within a nucleic acid-containing sample, the method comprising contacting the sample under suitable hybridization conditions with a nucleic acid probe of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The terms "complementary" or "complementarity" as used herein refer to the capacity of purine and pyrimidine nucleotides to associate through hydrogen bonding to form double stranded nucleic acid compounds. The following base pairs are related by complementarity: guanine and cytosine; adenine and thymine; and adenine and uracil. As used herein, "complementary" means that the aforementioned relationship applies to substantially all base pairs comprising two single-stranded nucleic acid compounds over the entire length of said molecules. "Partially complementary" refers to the aforementioned relationship in which one of two

single-stranded nucleic acid compounds is shorter in length than the other such that a portion of one of the molecules remains single-stranded.

"Conservative substitution" or "conservative amino acid substitution" refers to a replacement of one or more amino acid residue(s) in a protein or peptide as stipulated in Table 1.

"Fragment thereof" refers to a fragment, piece, or sub-region of a nucleic acid or protein molecule whose sequence is disclosed herein, such that the fragment comprises 5 or more amino acids, or 10 or more nucleotides that are contiguous in the parent protein or nucleic acid compound. Fragment thereof may or may not retain biological activity. For example, a fragment of a protein disclosed herein could be used as an antigen to raise a specific antibody against the parent protein molecule. When referring to a nucleic acid compound, "fragment thereof" refers to 10 or more contiguous nucleotides, derived from the parent nucleic acid, and also, owing to the genetic code, to the complementary sequence. For example if the fragment entails the sequence 5'-AGCTAG-3', then "fragment thereof" would also include the complementary sequence, 3'-TCGATC-5'.

The term "fusion protein" denotes a hybrid protein molecule not found in nature comprising a translational fusion or enzymatic fusion in which two or more different proteins or fragments thereof are covalently linked on a single polypeptide chain.

"Functional fragment" or "functionally equivalent fragment", as used herein, refers to a region, or fragment

of a full length protein, or sequence of amino acids that, for example, comprises an active site, or any other conserved motif, relating to biological function. Functional fragments are capable of providing a biological activity substantially similar to a full length protein disclosed herein, namely the ability to inhibit differentiation of bone marrow stem cells into osteoclasts. Functional fragments may be produced by cloning technology, or as the natural products of alternative splicing mechanisms.

"Host cell" refers to any eukaryotic or prokaryotic cell that is suitable for propagating and/or expressing a cloned gene contained on a vector that is introduced into said host cell by, for example, transformation or transfection, or the like.

TNFRsol refers to a nucleic acid and a protein or amino acid sequence encoded thereby. TNFRsol is a member of the TNFR superfamily. This family of receptors mediates a variety of biological effects of TNF ligands, including inhibition of bone resorption (by virtue of inhibiting osteoclast differentiation).

The term "homolog" or "homologous" describes the relationship between different nucleic acid compounds or amino acid sequences in which said sequences or molecules are related by partial identity or similarity at one or more blocks or regions within said molecules or sequences.

The term "hybridization" as used herein refers to a process in which a single-stranded nucleic acid compound joins with a complementary strand through nucleotide base pairing. The degree of hybridization depends upon, for

example, the degree of homology, the stringency of hybridization, and the length of hybridizing strands. "Selective hybridization" refers to hybridization under conditions of high stringency.

"Isolated nucleic acid compound" refers to any RNA or DNA sequence, however constructed or synthesized, which is locationally distinct from its natural location.

A "nucleic acid probe" or "probe" as used herein is a labeled nucleic acid compound which hybridizes with another nucleic acid compound. "Nucleic acid probe" means a single stranded nucleic acid sequence that will combine with a complementary or partially complementary single stranded target nucleic acid sequence to form a double-stranded molecule. A nucleic acid probe may be an oligonucleotide or a nucleotide polymer. A probe will usually contain a detectable moiety which may be attached to the end(s) of the probe or be internal to the sequence of the probe.

The term "orthologue" or "orthologous" refers to two or more genes or proteins from different organisms that exhibit sequence homology.

The term "paralogue" or "paralogous" refers to two or more genes or proteins within a single organism that exhibit sequence homology.

The term "plasmid" refers to an extrachromosomal genetic element. The plasmids disclosed herein are commercially available, publicly available on an unrestricted basis, or can be constructed from readily available plasmids in accordance with published procedures.

A "primer" is a nucleic acid fragment which functions as an initiating substrate for enzymatic or

synthetic elongation of, for example, a nucleic acid compound.

The term "promoter" refers to a nucleic acid sequence that directs transcription, for example, of DNA to RNA. An inducible promoter is one that is regulatable by environmental signals, such as carbon source, heat, or metal ions, for example. A constitutive promoter generally operates at a constant level and is not regulatable.

"Recombinant DNA cloning vector" as used herein refers to any autonomously replicating agent, including, but not limited to, plasmids and phages, comprising a DNA molecule to which one or more additional DNA segments can or have been incorporated.

The term "recombinant DNA expression vector" or "expression vector" as used herein refers to any recombinant DNA cloning vector, for example a plasmid or phage, in which a promoter and other regulatory elements are present thereby enabling transcription of an inserted DNA, which may encode a protein.

The term "stringency" refers to hybridization conditions. High stringency conditions disfavor non-homologous basepairing. Low stringency conditions have the opposite effect. Stringency may be altered, for example, by temperature and salt concentration.

"Low stringency" conditions comprise, for example, a temperature of about 37° C or less, a formamide concentration of less than about 50%, and a moderate to low salt (SSC) concentration; or, alternatively, a temperature of about 50° C or less, and a moderate to high salt (SSPE) concentration, for example 1M NaCl.

"High stringency" conditions comprise, for example, a temperature of about 42° C or less, a formamide concentration of less than about 20%, and a low salt (SSC) concentration; or, alternatively, a temperature of about 65° C, or less, and a low salt (SSPE) concentration. For example, high stringency conditions comprise hybridization in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C (Ausubel, F.M. *et al.* Current Protocols in Molecular Biology, Vol. I, 1989; Green Inc. New York, at 2.10.3).

"SSC" comprises a hybridization and wash solution. A stock 20X SSC solution contains 3M sodium chloride, 0.3M sodium citrate, pH 7.0.

"SSPE" comprises a hybridization and wash solution. A 1X SSPE solution contains 180 mM NaCl, 9 mM Na_2HPO_4, 0.9 mM NaH_2PO_4 and 1 mM EDTA, pH 7.4.

"Substantially pure" used in reference to a peptide or protein means that said peptide or protein is separated from other cellular and non-cellular molecules, including other protein molecules. A substantially pure preparation would be about at least 85% pure; preferably about at least 95% pure. A "substantially pure" protein as described herein could be prepared by a variety of techniques well known to the skilled artisan, including, for example, the IMAC protein purification method.

The term "vector" as used herein refers to a nucleic acid compound used for introducing exogenous or endogenous DNA into host cells. A vector comprises a nucleotide sequence which may encode one or more protein molecules. Plasmids, cosmids, viruses, and bacteriophages,

in the natural state or which have undergone recombinant engineering, are examples of commonly used vectors.

The various restriction enzymes disclosed and described herein are commercially available and the manner of use of said enzymes including reaction conditions, cofactors, and other requirements for activity are well known to one of ordinary skill in the art. Reaction conditions for particular enzymes were carried out according to the manufacturer's recommendation.

The TNFRsol gene comprises a nucleotide sequence of 900 nucleotide base pairs (SEQ ID NO:1) that encodes a polypeptide of 300 amino acid residues in length (SEQ ID NO:2). The TNFRsol gene identified from colon cells has a 87 nucleotide base pair sequence at the 5' end (i.e., nucleotides 1-87 of SEQ ID NO:1) that encodes a 29 residue signal peptide (i.e., residues 1-29 of SEQ ID NO:2), which peptide is cleaved from the N-terminus upon secretion of the mature soluble protein (i.e., residues 30-300 of SEQ ID NO:2).

Those skilled in the art will recognize that owing to the degeneracy of the genetic code (i.e. 64 codons which encode 20 amino acids), numerous "silent" substitutions of nucleotide base pairs could be introduced into the sequence identified as SEQ ID NO:1 without altering the identity of the encoded amino acid(s) or protein product. All such substitutions are intended to be within the scope of the invention.

Also contemplated by the present invention are TNFRsol proteins and related functional fragments such as, for example, smaller alternatively spliced forms, or

substitutions in which the primary sequence disclosed in SEQ ID NO:2 is altered by substitution or replacement or deletion or insertion at one or more amino acid positions, such that biological function is maintained. Functional fragments are conveniently identified as fragments of an intact TNFRsol protein that retain the capacity to inhibit osteoclast differentiation.

Several structural motifs have been identified within the primary sequence of TNFRsol protein that are thought to be important for biological function. For example, four cysteine rich motifs in the N-terminal domain, which are represented in a variety of related proteins, and which can form internal disulfide bonds, span from amino acid residue 34 to 195 of SEQ ID NO:2. It is presumed that this moiety retains biological function.

Functional analogs of the TNFRsol protein(s) are typically generated by deletion, insertion, or substitution of a single (or few) amino acid residues. Substitution modifications can generally be made in accordance with the following Table.

Table 1

ORIGINAL RESIDUE	EXEMPLARY SUBSTITUTIONS
ALA	SER
ARG	LYS
ASN	GLN, HIS
ASP	GLU
CYS	SER
GLN	ASN
GLU	ASP
GLY	PRO
HIS	ASN, GLN
ILE	LEU, VAL
LEU	ILE, VAL
LYS	ARG, GLN, GLU
MET	LEU, ILE
PHE	MET, LEU, TYR
SER	THR
THR	SER
TRP	TYR
TYR	TRP, PHE
VAL	ILE, LEU

Fragments of proteins

One embodiment of the instant invention provides fragments of the proteins disclosed that may or may not be biologically active. Such fragments are useful, for

example, as an antigen for producing an antibody to said proteins.

Fragments of the proteins disclosed herein may be generated by any number of suitable techniques, including chemical synthesis of any portion of SEQ ID NO:2, proteolytic digestion of SEQ ID NO:2, or most preferably, by recombinant DNA mutagenesis techniques, well known to the skilled artisan. *See. e.g.* K. Struhl, "Reverse biochemistry: Methods and applications for synthesizing yeast proteins *in vitro*," *Meth. Enzymol.* 194:520-535. For example, in a preferred method, a nested set of deletion mutations are introduced into the intact gene encoding the native TNFRsol protein such that varying amounts of the protein coding region are deleted, either from the amino terminal end, or from the carboxyl end of the protein molecule. This method can also be used to create internal fragments of the intact protein in which both the carboxyl and amino terminal ends are removed. Several appropriate nucleases can be used to create such deletions, for example *Bal*31, or in the case of a single stranded nucleic acid compound, mung bean nuclease. For simplicity, it is preferred that the intact TNFRsol gene be cloned into a single-stranded cloning vector, such as bacteriophage M13, or equivalent. If desired, the resulting gene deletion fragments can be subcloned into any suitable vector for propagation and expression of said fragments in any suitable host cell.

The present invention also provides fragments of the proteins disclosed herein wherein said fragments retain biological activity. As used herein, "functional fragments"

includes fragments of SEQ ID NO:2 that retain and exhibit, under appropriate conditions, measurable biological activity, for example, the capacity to inhibit osteoclast differentiation.

Functional fragments of the proteins disclosed herein may be produced as described above, preferably using cloning techniques to engineer smaller versions of the intact gene, lacking sequence from the 5' end, the 3' end, from both ends, or from an internal site.

Gene Isolation Procedures

Those skilled in the art will recognize that the TNFRsol gene could be obtained by a plurality of recombinant DNA techniques including, for example, hybridization, polymerase chain reaction (PCR) amplification, or de novo DNA synthesis. (*See e.g.*, T. Maniatis *et al.* Molecular Cloning: A Laboratory Manual, 2d Ed. Chap. 14 (1989)).

Methods for constructing cDNA libraries in a suitable vector such as a plasmid or phage for propagation in prokaryotic or eukaryotic cells are well known to those skilled in the art. (*See e.g.* Maniatis *et al. Supra*). Suitable cloning vectors are well known and are widely available.

The TNFRsol gene or fragment thereof can be isolated from any tissue in which said gene is expressed. In one method, mRNA is isolated from a suitable tissue, and first strand cDNA synthesis is carried out. A second round of DNA synthesis can be carried out for the production of the second strand. If desired, the double-stranded cDNA can be cloned into any suitable vector, for example, a plasmid,

thereby forming a cDNA library. Oligonucleotide primers targeted to any suitable region of SEQ ID NO:1 can be used for PCR amplification of TNFRsol. *See e.g.* PCR Protocols: A Guide to Method and Application, Ed. M. Innis et al., Academic Press (1990). The PCR amplification comprises template DNA, suitable enzymes, primers, and buffers, and is conveniently carried out in a DNA Thermal Cycler (Perkin Elmer Cetus, Norwalk, CT). A positive result is determined by detecting an appropriately-sized DNA fragment following agarose gel electrophoresis.

Protein Production Methods

One embodiment of the present invention relates to the substantially purified protein encoded by the TNFRsol gene.

Skilled artisans will recognize that the proteins of the present invention can be synthesized by a number of different methods, such as chemical methods well known in the art, including solid phase peptide synthesis or recombinant methods. Both methods are described in U.S. Patent 4,617,149, incorporated herein by reference.

The principles of solid phase chemical synthesis of polypeptides are well known in the art and may be found in general texts in the area. *See, e.g.*, H. Dugas and C. Penney, Bioorganic Chemistry (1981) Springer-Verlag, New York, 54-92. For example, peptides may be synthesized by solid-phase methodology utilizing an Applied Biosystems 430A peptide synthesizer (Applied Biosystems, Foster City, CA) and synthesis cycles supplied by Applied Biosystems.

The proteins of the present invention can also be produced by recombinant DNA methods using the cloned TNFRsol gene. Recombinant methods are preferred if a high yield is desired. Expression of the cloned gene can be carried out in a variety of suitable host cells, well known to those skilled in the art. For this purpose, the TNFRsol gene is introduced into a host cell by any suitable means, well known to those skilled in the art. While chromosomal integration of the cloned gene is within the scope of the present invention, it is preferred that the gene be cloned into a suitable extra-chromosomally maintained expression vector so that the coding region of the TNFRsol gene is operably-linked to a constitutive or inducible promoter.

The basic steps in the recombinant production of the TNFRsol protein are:

a) constructing a natural, synthetic or semi-synthetic DNA encoding TNFRsol protein;

b) integrating said DNA into an expression vector in a manner suitable for expressing the TNFRsol protein, either alone or as a fusion protein;

c) transforming or otherwise introducing said vector into an appropriate eukaryotic or prokaryotic host cell forming a recombinant host cell;

d) culturing said recombinant host cell in a manner to express the TNFRsol protein; and

e) recovering and substantially purifying the TNFRsol protein by any suitable means well known to those skilled in the art.

Expressing Recombinant TNFRsol Protein in Prokaryotic and Eukaryotic Host Cells

Prokaryotes may be employed in the production of recombinant TNFRsol protein. For example, the *Escherichia coli* K12 strain 294 (ATCC No. 31446) is particularly useful for the prokaryotic expression of foreign proteins. Other strains of *E. coli*, bacilli such as *Bacillus subtilis*, enterobacteriaceae such as *Salmonella typhimurium* or *Serratia marcescans*, various Pseudomonas species and other bacteria, such as *Streptomyces*, may also be employed as host cells in the cloning and expression of the recombinant proteins of this invention.

Promoter sequences suitable for driving the expression of genes in prokaryotes include β-lactamase [*e.g.* vector pGX2907, ATCC 39344, contains a replicon and β-lactamase gene], lactose systems [Chang et al., *Nature* (London), 275:615 (1978); Goeddel et al., *Nature* (London), 281:544 (1979)], alkaline phosphatase, and the tryptophan (trp) promoter system [vector pATH1 (ATCC 37695)], which is designed to facilitate expression of an open reading frame as a trpE fusion protein under the control of the trp promoter. Hybrid promoters such as the tac promoter (isolatable from plasmid pDR540, ATCC-37282) are also suitable. Still other bacterial promoters, whose nucleotide sequences are generally known, may be ligated to DNA encoding the protein of the instant invention, using linkers or adapters to supply any required restriction sites. Promoters for use in bacterial systems also will contain a Shine-Dalgarno sequence operably-linked to the DNA encoding

the desired polypeptides. These examples are illustrative rather than limiting.

The proteins of this invention may be synthesized either by direct expression or as a fusion protein comprising the protein of interest as a translational fusion with another protein or peptide which may be removed by enzymatic or chemical cleavage. It is often observed in the production of certain peptides in recombinant systems that expression as a fusion protein prolongs the lifespan, increases the yield of the desired peptide, or provides a convenient means of purifying the protein. This is particularly relevant when expressing mammalian proteins in prokaryotic hosts. A variety of peptidases (e.g. enterokinase and thrombin) which cleave a polypeptide at specific sites or digest the peptides from the amino or carboxy termini (*e.g.* diaminopeptidase) of the peptide chain are known. Furthermore, particular chemicals (*e.g.* cyanogen bromide) will cleave a polypeptide chain at specific sites. The skilled artisan will appreciate the modifications necessary to the amino acid sequence (and synthetic or semi-synthetic coding sequence if recombinant means are employed) to incorporate site-specific internal cleavage sites. *See e.g.*, P. Carter, "Site Specific Proteolysis of Fusion Proteins", Chapter 13, in Protein Purification: From Molecular Mechanisms to Large Scale Processes, American Chemical Society, Washington, D.C. (1990).

In addition to prokaryotes, a variety of amphibian expression systems such as frog oocytes, and mammalian cell systems can be used. The choice of a particular host cell depends to some extent on the particular expression vector

used. Exemplary mammalian host cells suitable for use in the present invention include HepG-2 (ATCC HB 8065), CV-1 (ATCC CCL 70), LC-MK_2 (ATCC CCL 7.1), 3T3 (ATCC CCL 92), CHO-K1 (ATCC CCL 61), HeLa (ATCC CCL 2), RPMI8226 (ATCC CCL 155), H4IIEC3 (ATCC CCL 1600), C127I (ATCC CCL 1616), HS-Sultan (ATCC CCL 1484), and BHK-21 (ATCC CCL 10), for example.

A wide variety of vectors are suitable for transforming mammalian host cells. For example, the pSV2-type vectors comprise segments of the simian virus 40 (SV40) genome required for transcription and polyadenylation. A large number of plasmid pSV2-type vectors have been constructed, such as pSV2-gpt, pSV2-neo, pSV2-dhfr, pSV2-hyg, and pSV2-β-globin, in which the SV40 promoter drives transcription of an inserted gene. These vectors are widely available from sources such as the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland, 20852, or the National Center for Agricultural Utilization Research, 1815 North University Street, Peoria, Illinois 61604-39999.

Promoters suitable for expression in mammalian cells include the SV40 late promoter, promoters from eukaryotic genes, such as, for example, the estrogen-inducible chicken ovalbumin gene, the interferon genes, the glucocorticoid-inducible tyrosine aminotransferase gene, the thymidine kinase gene promoter, and the promoters of the major early and late adenovirus genes.

Plasmid pRSVcat (ATCC 37152) comprises portions of a long terminal repeat of the Rous Sarcoma virus, a virus known to infect chickens and other host cells. This long

terminal repeat contains a promoter which is suitable for use in the vectors of this invention. H. Gorman *et al.*, *Proc. Nat. Acad. Sci. (USA)*, 79, 6777 (1982). The plasmid pMSVi (NRRL B-15929) comprises the long terminal repeats of the Murine Sarcoma virus, a virus known to infect mouse and other host cells. The mouse metallothionein promoter has also been well characterized for use in eukaryotic host cells and is suitable for use in the present invention. This promoter is present in the plasmid pdBPV-MMTneo (ATCC 37224) which can serve as the starting material for the construction of other plasmids of the present invention.

Transfection of mammalian cells with vectors can be performed by a plurality of well known processes including, but not limited to, protoplast fusion, calcium phosphate co-precipitation, electroporation and the like. *See, e.g.*, Maniatis *et al.*, *supra*.

Some viruses also make appropriate vectors. Examples include the adenoviruses, the adeno-associated viruses, the vaccinia virus, the herpes viruses, the baculoviruses, and the rous sarcoma virus, as described in U.S. Patent 4,775,624, incorporated herein by reference.

Eukaryotic microorganisms such as yeast and other fungi are also suitable host cells. The yeast *Saccharomyces cerevisiae* is the preferred eukaryotic microorganism. Other yeasts such as *Kluyveromyces lactis* and *Pichia pastoris* are also suitable. For expression in *Saccharomyces*, the plasmid YRp7 (ATCC-40053), for example, may be used. *See, e.g.*, L. Stinchcomb *et al.*, *Nature*, 282, 39 (1979); J. Kingsman *et al.*, *Gene*, 7, 141 (1979); S. Tschemper *et al.*, *Gene*, 10, 157

(1980). Plasmid YRp7 contains the TRP1 gene which provides a selectable marker for use in a trp1 auxotrophic mutant.

Purification of Recombinantly-Produced TNFRsol Protein

An expression vector carrying the cloned TNFRsol gene is transformed or transfected into a suitable host cell using standard methods. Cells that contain the vector are propagated under conditions suitable for expression of the recombinant TNFRsol protein. For example, if the recombinant gene has been placed under the control of an inducible promoter, suitable growth conditions would incorporate the appropriate inducer. The recombinantly-produced protein may be purified from cellular extracts of transformed cells by any suitable means.

In a preferred process for protein purification, the TNFRsol gene is modified at the 5' end to incorporate several histidine residues at the amino terminus of the TNFRsol protein. This "histidine tag" enables a single-step protein purification method referred to as "immobilized metal ion affinity chromatography" (IMAC), essentially as described in U.S. Patent 4,569,794, which hereby is incorporated by reference. The IMAC method enables rapid isolation of substantially pure recombinant TNFRsol protein starting from a crude extract of cells that express a modified recombinant protein, as described above.

Production of Antibodies

The proteins of this invention and fragments thereof may be used in the production of antibodies. The term "antibody" as used herein describes antibodies,

fragments of antibodies (such as, but not limited, to Fab, Fab', Fab_2', and Fv fragments), and chimeric, humanized, veneered, resurfaced, or CDR-grafted antibodies capable of binding antigens of a similar nature as the parent antibody molecule from which they are derived. The instant invention also encompasses single chain polypeptide binding molecules.

The production of antibodies, both monoclonal and polyclonal, in animals, especially mice, is well known in the art. *See, e.g.*, C. Milstein, Handbook of Experimental Immunology, (Blackwell Scientific Pub., 1986); J. Goding, Monoclonal Antibodies: Principles and Practice, (Academic Press, 1983). For the production of monoclonal antibodies the basic process begins with injecting a mouse, or other suitable animal, with an immunogen. The mouse is subsequently sacrificed and cells taken from its spleen are fused with myeloma cells, resulting in a hybridoma that reproduces *in vitro*. The population of hybridomas is screened to isolate individual clones, each of which secretes a single antibody species, specific for the immunogen. Each antibody obtained in this way is the clonal product of a single B cell.

Chimeric antibodies are described in U.S. Patent No. 4,816,567, the entire contents of which is herein incorporated by reference. This reference discloses methods and vectors for the preparation of chimeric antibodies. An alternative approach is provided in U.S. Patent No. 4,816,397, the entire contents of which is herein incorporated by reference. This patent teaches co-expression of the heavy and light chains of an antibody in the same host cell.

The approach of U.S. Patent 4,816,397 has been further refined in European Patent Publication No. 0 239 400. The teachings of this European patent publication are a preferred format for genetic engineering of monoclonal antibodies. In this technology the complementarity determining regions (CDRs) of a human antibody are replaced with the CDRs of a murine monoclonal antibody, thereby converting the specificity of the human antibody to the specificity of a murine antibody.

Single chain antibodies and libraries thereof are yet another variety of genetically engineered antibody technology that is well known in the art. (*See, e.g.* R.E. Bird, *et al.*, *Science* 242:423-426 (1988); PCT Publication Nos. WO 88/01649, WO 90/14430, and WO 91/10737. Single chain antibody technology involves covalently joining the binding regions of heavy and light chains to generate a single polypeptide chain. The binding specificity of the intact antibody molecule is thereby reproduced on a single polypeptide chain.

The antibodies contemplated by this invention are useful in diagnostics, therapeutics or in diagnostic/therapeutic combinations.

The proteins of this invention or suitable fragments thereof can be used to generate polyclonal or monoclonal antibodies, and various inter-species hybrids, or humanized antibodies, or antibody fragments, or single-chain antibodies. The techniques for producing antibodies are well known to skilled artisans. (*See e.g.* A.M. Campbell, Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science

Publishers, Amsterdam (1984); Kohler and Milstein, *Nature* 256, 495-497 (1975); Monoclonal Antibodies: Principles & Applications Ed. J.R.Birch & E.S. Lennox, Wiley-Liss, 1995.

A protein used as an immunogen may be modified or administered in an adjuvant, by subcutaneous or intraperitoneal injection into, for example, a mouse or a rabbit. For the production of monoclonal antibodies, spleen cells from immunized animals are removed, fused with myeloma cells, such as SP2/0-Ag14 cells, and allowed to become monoclonal antibody producing hybridoma cells in the manner known to the skilled artisan. Hybridomas that secrete a desired antibody molecule can be screened by a variety of well known methods, for example ELISA assay, western blot analysis, or radioimmunoassay (Lutz *et al. Exp. Cell Res.* 175, 109-124 (1988); Monoclonal Antibodies: Principles & Applications Ed. J.R.Birch & E.S. Lennox, Wiley-Liss, 1995).

For some applications labeled antibodies are desirable. Procedures for labeling antibody molecules are widely known, including for example, the use of radioisotopes, affinity labels, such as biotin or avidin, enzymatic labels, for example horseradish peroxidase, and fluorescent labels, such as FITC or rhodamine (*See e.g.* Enzyme-Mediated Immunoassay, Ed. T. Ngo, H. Lenhoff, Plenum Press 1985; Principles of Immunology and Immunodiagnostics, R.M. Aloisi, Lea & Febiger, 1988).

Labeled antibodies are useful for a variety of diagnostic applications. In one embodiment the present invention relates to the use of labeled antibodies to detect the presence of TNFRsol. Alternatively, the antibodies could be used in a screen to identify potential modulators

of TNFRsol. For example, in a competitive displacement assay, the antibody or compound to be tested is labeled by any suitable method. Competitive displacement of an antibody from an antibody-antigen complex by a test compound such that a test compound-antigen complex is formed provides a method for identifying compounds that bind HPLFP.

Other embodiments of the present invention comprise isolated nucleic acid sequences that encode SEQ ID NO:2, or fragments thereof. As skilled artisans will recognize, the amino acid compounds of the invention can be encoded by a multitude of different nucleic acid sequences, owing to the degeneracy of the genetic code. Because these alternative nucleic acid sequences would encode the same amino acid sequences, the present invention further comprises these alternate nucleic acid sequences. Also contemplated are related nucleic acids that are at least about 75% identical to SEQ ID NO:1, or to their complementary sequence, or nucleic acids that hybridize to SEQ ID NO:1 under low stringency conditions. Such sequences may come, for example, from other related genes.

The TNFRsol gene (*viz.* SEQ ID NO:1) and related nucleic acid compounds that encode SEQ ID NO:2, or functional fragments thereof, may be produced by chemical synthetic methods. The synthesis of nucleic acids is well known in the art. *See, e.g.*, E.L. Brown, R. Belagaje, M.J. Ryan, and H.G. Khorana, *Methods in Enzymology*, 68:109-151 (1979). Fragments of the DNA sequence corresponding to the TNFRsol gene could be generated using a conventional DNA synthesizing apparatus, such as the Applied Biosystems Model 380A or 380B DNA synthesizers (Applied Biosystems, Inc., 850

Lincoln Center Drive, Foster City, CA 94404) using phosphoramidite chemistry, thereafter ligating the fragments so as to reconstitute the entire gene. Alternatively, phosphotriester chemistry may be employed to synthesize the nucleic acids of this invention. (*See, e.g.*, M.J. Gait, ed., Oligonucleotide Synthesis, A Practical Approach, (1984)).

In an alternative methodology, namely PCR, the DNA sequences disclosed and described herein, comprising, for example, a portion or all of SEQ ID NO:1 can be produced from a plurality of starting materials. For example, starting with a cDNA preparation (e.g. cDNA library) derived from a tissue that expresses the TNFRsol gene, suitable oligonucleotide primers complementary to SEQ ID NO:1 or to any sub-region therein, are prepared as described in U.S. Patent No. 4,889,818, hereby incorporated by reference. Other suitable protocols for the PCR are disclosed in PCR Protocols: A Guide to Method and Applications, Ed. Michael A. Innis et al., Academic Press, Inc. (1990). Using PCR, any region of the TNFRsol gene can be targeted for amplification such that full or partial length gene sequences may be produced.

The ribonucleic acids of the present invention may be prepared using polynucleotide synthetic methods discussed *supra*, or they may be prepared enzymatically, for example, using RNA polymerase to transcribe a TNFRsol DNA template.

The most preferred systems for preparing the ribonucleic acids of the present invention employ the RNA polymerase from the bacteriophage T7 or the bacteriophage SP6. These RNA polymerases are highly specific, requiring

the insertion of bacteriophage-specific sequences at the 5' end of the template to be transcribed. *See*, Maniatis *et al.*, *supra*.

This invention also provides nucleic acids, RNA or DNA, that are complementary to SEQ ID NO:1, or fragments thereof.

Nucleic Acid Probes

The present invention also provides probes and primers useful for a variety of molecular biology techniques including, for example, hybridization screens of genomic, subgenomic, or cDNA libraries, as well as hybridization against nucleic acids derived from cell lines or tissues that originate from drug-resistant tumors. Such hybridization screens are useful as methods to identify homologous and/or functionally related sequences from the same or other organisms, and further for investigating the mechanism by which drug resistance arises in various cancers. A nucleic acid compound comprising SEQ ID NO:1 or a complementary sequence thereof, or fragment thereof, which is at least 14 base pairs in length, and which will selectively hybridize to human DNA or mRNA encoding TNFRsol protein or fragment thereof, or a functionally related protein, is provided. Preferably, the 14 or more base pair compound is DNA. *See e.g.* B. Wallace and G. Miyada, "Oligonucleotide Probes for the Screening of Recombinant DNA Libraries," In *Meth. Enzym.*, 152, 432-442, Academic Press (1987).

Probes and primers can be prepared by enzymatic or recombinant methods, well known to those skilled in the art

(*See e.g.* Sambrook et al. *supra*). A probe may be a single stranded nucleic acid sequence which is complementary in some particular degree to a nucleic acid sequence sought to be detected ("target sequence"). A probe may be labeled with a detectable moiety such as a radio-isotope, antigen, or chemiluminescent moiety. A description of the use of nucleic acid hybridization as a procedure for the detection of particular nucleic acid sequences is described in U.S. Patent No. 4,851,330 to Kohne, entitled "Method for Detection, Identification and Quantitation of Non-Viral Organisms."

DNA sequence information provided by the present invention allows for the preparation of relatively short DNA (or RNA) sequences having the ability to specifically hybridize to gene sequences disclosed herein. In these aspects, nucleic acid probes of an appropriate length are prepared. The ability of such nucleic acid probes to specifically hybridize to a polynucleotide encoding a TNFRsol gene or related sequence lends particular utility in a variety of embodiments. Most importantly, the probes may be used in a variety of assays for detecting the presence of complementary sequences in a given sample.

In certain embodiments, it is advantageous to use oligonucleotide primers. The sequence of such primers is designed using a polynucleotide of the present invention for use in detecting, amplifying, or mutating a defined segment of a gene or polynucleotide that encodes a TNFRsol polypeptide using PCR technology.

Preferred nucleic acid sequences employed for hybridization studies, or assays, include probe molecules

that are complementary to at least an about 14- to an about 70-nucleotide long stretch of a polynucleotide that encodes a TNFRsol polypeptide, such as the nucleotide base sequences designated as SEQ ID NO:1. A length of at least 14 nucleotides helps to ensure that the fragment is of sufficient length to form a duplex molecule that is both stable and selective. Molecules having complementary sequences over stretches greater than 14 bases in length are generally preferred in order to increase stability and selectivity of the hybrid. One will generally prefer to design nucleic acid compounds having gene-complementary stretches of 25 to 40 nucleotides, 55 to 70 nucleotides, or even longer where desired. Such fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, by application of nucleic acid reproduction technology, such as the PCR TM technology of U.S. Pat. No. 4,603,102, herein incorporated by reference, or by excising selected DNA fragments from recombinant plasmids containing appropriate inserts and suitable restriction enzyme sites.

The following guidelines are useful for designing probes with desirable characteristics. The extent and specificity of hybridization reactions are affected by a number of factors that determine the sensitivity and specificity of a particular probe, whether perfectly complementary to its target or not. The affect of various experimental parameters and conditions are well known to those skilled in the art.

First, the stability of the probe:target nucleic acid hybrid should be chosen to be compatible with the assay

conditions. This may be accomplished by avoiding long A and T rich sequences, by terminating the hybrids with G:C base pairs and by designing a probe with an appropriate Tm (i.e. melting temperature). The melting profile, including the Tm of a hybrid comprising an oligonucleotide and target sequence, may be determined using a Hybridization Protection Assay. The probe should be chosen so that the length and percent GC content result in a Tm about 2°-10° C higher than the temperature at which the final assay will be performed. The base composition of the probe is also a significant factor because G-C base pairs exhibit greater thermal stability as compared to A-T base pairs. Thus, hybridization involving complementary nucleic acids of higher G-C content will be more stable at higher temperatures.

The ionic strength and incubation temperature under which a probe will be used should also be taken into account. It is known that hybridization will increase as the ionic strength of the reaction mixture increases, and that the thermal stability of molecular hybrids will increase with increasing ionic strength. On the other hand, chemical reagents such as formamide, urea, DMSO and alcohols, which disrupt hydrogen bonds, increase the stringency of hybridization. Destabilization of hydrogen bonds by such reagents can greatly reduce the Tm. In general, optimal hybridization for synthetic oligonucleotide probes of about 10-50 bases in length occurs approximately 5° C below the melting temperature for a given duplex. Incubation at temperatures below the optimum may allow

mismatched base sequences to hybridize and can therefore result in reduced specificity.

The length of the target nucleic acid sequence and, accordingly, the length of the probe sequence can also be important. In some cases, there may be several sequences from a particular region, varying in location and length, which will yield probes with the desired hybridization characteristics. In other cases, one sequence may be significantly better than another even though the one sequence differs merely by a single base. Finally, there can be intramolecular and intermolecular hybrids formed within a probe if there is sufficient self-complementarity. Such structures can be avoided through careful probe design. Computer programs are available to search for this type of interaction.

A probe molecule may be used for hybridizing to a sample suspected of possessing a TNFRsol or TNFRsol-related nucleotide sequence. The hybridization reaction is carried out under suitable conditions of stringency.

Alternatively, such DNA molecules may be used in a number of techniques including their use as: (1) diagnostic tools to detect polymorphisms in DNA samples from a human or other mammal; (2) means for detecting and isolating homologs of TNFRsol and related polypeptides from a DNA library potentially containing such sequences; (3) primers for hybridizing to related sequences for the purpose of amplifying those sequences; and (4) primers for altering the native TNFRsol DNA sequences; as well as other techniques which rely on the similarity of the DNA sequences to those of the TNFRsol DNA segments herein disclosed.

Once synthesized, oligonucleotide probes may be labeled by any of several well known methods. *See e.g.* Maniatis *et.al.*, Molecular Cloning (2d ed. 1989). Useful labels include radioisotopes, as well as non-radioactive reporting groups. Isotopic labels include H^3, S^{35}, P^{32}, I^{125}, Cobalt, and C^{14}. Most methods of isotopic labeling involve the use of enzymes and include methods such as nick-translation, end-labeling, second strand synthesis, and reverse transcription. When using radio-labeled probes, hybridization can be detected by autoradiography, scintillation counting, or gamma counting. The detection method selected will depend upon the hybridization conditions and the particular radio isotope used for labeling.

Non-isotopic materials can also be used for labeling, and may be introduced internally into the sequence or at the end of the sequence. Modified nucleotides may be incorporated enzymatically or chemically, and chemical modifications of the probe may be performed during or after synthesis of the probe, for example, by the use of non-nucleotide linker groups. Non-isotopic labels include fluorescent molecules, chemiluminescent molecules, enzymes, cofactors, enzyme substrates, haptens or other ligands.

In a preferred embodiment of the invention, the length of an oligonucleotide probe is greater than or equal to about 18 nucleotides and less than or equal to about 50 nucleotides. Labeling of an oligonucleotide of the present invention may be performed enzymatically using [^{32}P]-labeled ATP and the enzyme T4 polynucleotide kinase.

Vectors

Another aspect of the present invention relates to recombinant DNA cloning vectors and expression vectors comprising the nucleic acids of the present invention. The preferred nucleic acid vectors are those which comprise DNA, in particular SEQ ID NO:1, more particularly nucleotides 88-900 of SEQ ID NO:1, and more particularly nucleotides 102-585 of SEQ ID NO:1.

The skilled artisan understands that choosing the most appropriate cloning vector or expression vector depends upon a number of factors including the availability of restriction enzyme sites, the type of host cell into which the vector is to be transfected or transformed, the purpose of the transfection or transformation (e.g., stable transformation as an extrachromosomal element, or integration into the host chromosome), the presence or absence of readily assayable or selectable markers (e.g., antibiotic resistance and metabolic markers of one type and another), and the number of copies of the gene desired in the host cell.

Vectors suitable to carry the nucleic acids of the present invention comprise RNA viruses, DNA viruses, lytic bacteriophages, lysogenic bacteriophages, stable bacteriophages, plasmids, viroids, and the like. The most preferred vectors are plasmids.

When preparing an expression vector the skilled artisan understands that there are many variables to be considered, for example, whether to use a constitutive or inducible promoter. The practitioner also understands that the amount of nucleic acid or protein to be produced

dictates, in part, the selection of the expression system. Regarding promoter sequences, inducible promoters are preferred because they enable high level, regulatable expression of an operably-linked gene. The skilled artisan will recognize a number of suitable promoters that respond to a variety of inducers, for example, carbon source, metal ions, and heat. Other relevant considerations regarding an expression vector include whether to include sequences for directing the localization of a recombinant protein. For example, a sequence encoding a signal peptide preceding the coding region of a gene is useful for directing the extracellular export of a resulting polypeptide.

The present invention also provides a method for constructing a recombinant host cell capable of expressing proteins comprising SEQ ID NO:2, said method comprising transforming or otherwise introducing into a host cell a recombinant DNA vector that comprises an isolated DNA sequence that encodes SEQ ID NO:2. Of course. Such method also encompasses the host cells capable of expressing functional fragments of SEQ ID NO:2. The preferred host cell is any eukaryotic cell that can accomodate high level expression of an exogenously introduced gene or protein, and that will incorporate said protein into its membrane structure. Vectors for expression are those which comprise SEQ ID NO:1 or a fragment thereof. Transformed host cells may be cultured under conditions well known to skilled artisans such that SEQ ID NO:2 is expressed, thereby producing a recombinant TNFRsol protein in the recombinant host cell.

For the purpose of identifying compounds having utility as regulators or modifiers of bone resorption, it would be desirable to identify compounds that bind the TNFRsol protein and/or modify its activity. A method for identifying such compounds comprises the steps of admixing a substantially purified preparation of a TNFRsol protein with a test compound, and monitoring by any suitable means a binding interaction between said protein and said compound.

Functional fragments of the proteins disclosed herein may also be identified as having activity. For this purpose, gene fragments (prepared as described elsewhere herein) are cloned into a suitable expression vector, and transformed or transfected into a suitable host cell. The culture medium of transformed or transfected host cells is then assayed for the ability to inhibit osteoclast differentiation. The level of activity in the transformed cells is compared to a negative control in which the organism is transformed by a vector without a TNFRsol insert and to a positive control in which the entire TNFRsol protein is present on the transforming vector. Fragments of TNFRsol that impart activity to about 30% or greater of the positive control cells are regarded as biologically functional.

Skilled artisans will recognize that IC_{50} values are dependent on the selectivity of the compound tested. For example, a compound with an IC_{50} which is less than 10 nM is generally considered an excellent candidate for drug therapy. However, a compound which has a lower affinity, but is selective for a particular target, may be an even better candidate. The skilled artisan will recognize that

any information regarding the binding potential, inhibitory activity, or selectivity of a particular compound is useful toward the development of pharmaceutical products.

The following examples more fully describe the present invention. Those skilled in the art will recognize that the particular reagents, equipment, and procedures described are merely illustrative and are not intended to limit the present invention in any manner.

EXAMPLE 1

RT-PCR Amplification of TNFRsol Gene from mRNA

A TNFRsol gene is isolated by reverse transcriptase PCR (RT-PCR) using conventional methods. Total RNA from a tissue that expresses the TNFRsol gene, for example, lung, is prepared using standard methods. First strand TNFRsol cDNA synthesis is achieved using a commercially available kit (SuperScript™ System; Life Technologies) by PCR in conjunction with specific primers directed at any suitable region of SEQ ID NO:1.

Amplification is carried out by adding to the first strand cDNA (dried under vacuum): 8 µl of 10X synthesis buffer (200 mM Tris-HCl, pH 8.4; 500 mM KCl, 25 mM $MgCl_2$, 1 ug/ul BSA); 68 µl distilled water; 1 µl each of a 10 uM solution of each primer; and 1 µl Taq DNA polymerase (2 to 5 U/µl). The reaction is heated at 94° C for 5 minutes to denature the RNA/cDNA hybrid. Then, 15 to 30 cycles of PCR amplification are performed using any suitable thermal cycle apparatus. The amplified sample may be analyzed by agarose gel electrophoresis to check for an appropriately-sized fragment.

EXAMPLE 2

Production of a Vector for Expressing TNFRsol in a Host Cell

An expression vector suitable for expressing TNFRsol or fragment thereof in a variety of prokaryotic host cells, such as *E. coli* is easily made. The vector contains an origin of replication (Ori), an ampicillin resistance gene (Amp) useful for selecting cells which have incorporated the vector following a transformation procedure, and further comprises the T7 promoter and T7 terminator sequences in operable linkage to a TNFRsol coding region. Plasmid pET11A (obtained from Novogen, Madison WI) is a suitable parent plasmid. pET11A is linearized by restriction with endonucleases NdeI and BamHI. Linearized pET11A is ligated to a DNA fragment bearing NdeI and BamHI sticky ends and comprising the coding region of the TNFRsol gene as disclosed by SEQ ID NO:1 or a fragment thereof.

The TNFRsol gene used in this construction may be slightly modified at the 5' end (amino terminus of encoded protein) in order to simplify purification of the encoded protein product. For this purpose, an oligonucleotide encoding 8 histidine residues is inserted after the ATG start codon. Placement of the histidine residues at the amino terminus of the encoded protein serves to enable the IMAC one-step protein purification procedure.

EXAMPLE 3

Recombinant Expression and Purification of TNFRsol Protein

An expression vector that carries an open reading frame (ORF) encoding TNFRsol or fragment thereof and which ORF is operably-linked to an expression promoter is transformed into *E. coli* BL21 (DE3) (*hsdS gal* λcIts857 *ind1Sam7nin5lac*UV5-T7gene 1) using standard methods. Transformants, selected for resistance to ampicillin, are chosen at random and tested for the presence of the vector by agarose gel electrophoresis using quick plasmid preparations. Colonies which contain the vector are grown in L broth and the protein product encoded by the vector-borne ORF is purified by immobilized metal ion affinity chromatography (IMAC), essentially as described in US Patent 4,569,794.

Briefly, the IMAC column is prepared as follows. A metal-free chelating resin (e.g. Sepharose 6B IDA, Pharmacia) is washed in distilled water to remove preservative substances and infused with a suitable metal ion [e.g. Ni(II), Co(II), or Cu(II)] by adding a 50mM metal chloride or metal sulfate aqueous solution until about 75% of the interstitial spaces of the resin are saturated with colored metal ion. The column is then ready to receive a crude cellular extract containing the recombinant protein product.

After removing unbound proteins and other materials by washing the column with any suitable buffer, pH 7.5, the bound protein is eluted in any suitable buffer at pH 4.3, or preferably with an imidizole-containing buffer at pH 7.5.

EXAMPLE 4

Tissue Distribution of TNFRsol mRNA

The presence of TNFRsol mRNA in a variety of human tissues was analyzed by Northern analysis. Total RNA from different tissues or cultured cells was isolated by a standard guanidine chloride/phenol extraction method, and poly-A$^+$ RNA was isolated using oligo(dT)-cellulose type 7 (Pharmacia). Electrophoresis of RNA samples was carried out in formaldehyde followed by capillary transfer to Zeta-Probe™ nylon membranes (Bio-Rad, Hercules, Calif.). SEQ ID NO:1 was the template for generating probes using a MultiPrime™ random priming kit (Amersham, Arlington Heights, Ill.). The efficiency of the labeling reaction was approximately 4 x 10^{10} cpm incorporated per µg of template. The hybridization buffer contained 0.5M sodium phosphate, 7% SDS (wt/vol), 1% BSA (wt/vol), and 1 mM EDTA. Prehybridization was carried out in hybridization buffer at 65° C for 2 h and ^{32}P-labeled probe was added and incubation continued overnight. The filters were washed in Buffer A (40 mM sodium phosphate pH 7.2, 5% SDS [wt/vol], 0.5% BSA [wt/vol], and 1 mM EDTA) at 65° C for 1 h, and then in Buffer B (40 mM sodium phosphate, pH 7.2, 1% SDS [wt/vol], and 1 mM EDTA) at 65° C for 20 minutes. The filters were air-dried and exposed to Kodak X-OMAT AR film at -80° C with an intensifying screen.

The results showed that TNFRsol mRNA was present in numerous tissues, including stomach, spinal cord, lymph node, trachea, spleen, and lung.

EXAMPLE 5

Production of an Antibody to a Protein

Substantially pure protein or fragment thereof is isolated from transfected or transformed cells using any of the well known methods in the art, or by a method specifically disclosed herein. Concentration of protein in a final preparation is adjusted, for example, by filtration through an Amicon filter device such that the level is about 1 to 5 ug/ml. Monoclonal or polyclonal antibody can be prepared as follows.

Monoclonal antibody can be prepared from murine hybridomas according to the method of Kohler and Milstein (*Nature*, 256, 495, 1975), or a modified method thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the protein or fragment thereof, or fusion peptide thereof, over a period of a few weeks. The mouse is then sacrificed and the antibody producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells. Fused cells that produce antibody are identified by any suitable immunoassay, for example, ELISA, as described in E. Engvall, *Meth. Enzymol.*, 70, 419, 1980.

Polyclonal antiserum can be prepared by well known methods (*See e.g.* J. Vaitukaitis *et.al. Clin. Endocirnol. Metab.* 33, 988, 1971) that involve immunizing suitable animals with the proteins, fragments thereof, or fusion proteins thereof, disclosed herein. Small doses (e.g. nanogram amounts) of antigen administered at multiple intradermal sites appears to be the most reliable method.

EXAMPLE 6

Murine Osteoclast Differentiation Assay

The co-culture method of Takahashi et al. (Endocrinology 123:2600 1988) was modified as described in Galvin et al. (Endocrinology 137:2457 1996) and used to study the effects of various agents on osteoclast differentiation.

Male Balb/C mice (4-8 weeks old) were euthanized with CO_2, the femurs removed, and the marrow flushed out of the femurs with growth medium. Bone marrow cells were pelleted by centrifugation at 500 x g for 6 min. and resuspended in the growth medium (RPMI 1640 plus 5% heat-inactivated fetal bovine-serum and 1% antibiotic-antimycotic solution). The marrow population (5×10^4 cells/cm^2) was seeded in tissue culture dishes in which BALC cells (a stable cell line derived from neonatal mouse calvariae, 1.5×10^4 cells/cm^2) had been plated 2 h prior to addition of bone marrow. The cells were cultured for 7 days in a humidified incubator at 37°C with 5% CO_2 with medium changes on days 3 and 5. Cultures were treated with or without 10^{-8}M $1,25\text{-}(OH)_2D_3$ on days 0, 3, and 5. In addition, the cells were treated with or without secreted TNFRsol protein purified from the conditioned medium of cells transfected with TNFRsol gene (SEQ ID NO:1). Following 7 days of culture, the cells in 24-well cluster dishes were fixed with formalin (3.7% for 10 min) and then stained for tartrate-resistant acid phosphatase (TRAP) using a modification of the method described by Graves, L and Jilka RL, J Cell Physiology 145:102 1990. The number of osteoclasts (TRAP-

positive cells containing 3 or more nuclei) was quantitated. Results are reported in Table 1.

Table 1

TNFRsol (ng/ml)	Osteoclasts/well[a]
0.00	145.50 ± 7.33
0.01	40.50 ± 2.39*
0.10	65.50 ± 3.33*
1.00	97.50 ± 3.10*
10.00	170.17 ± 8.26
100.00	335.00 ± 8.90*

a - Each value represents the mean and standard error of 6 wells.

*$p<0.05$ compared to control group

EXAMPLE 7

Porcine Osteoclast Differentiation Assay

Neonatal pigs (aged 1-5 days) were euthanized with CO_2, the appendages were rinsed with 70% ethanol, the soft tissues were removed, and the humeri, radii, ulnae, femora, tibiae and fibulae were excised. The long bones were placed in ice-cold calcium and magnesium-free Hank's balanced salt solution (CMF-HBSS, Gibco BRL) and cleaned of all soft tissues. The bones were split longitudinally and the endosteal surfaces were scraped to remove both the marrow and trabecular bone. The suspension of trabecular bone

particles and marrow cells was agitated by vigorous shaking and passed through a 200 mm and then 100 mm sieve. Cells were centrifuged at 500 x g for 10 minutes at 4°C, the pellet was resuspended in CMF-HBSS, and then separated on a Ficoll-Paque gradient (Pharmacia, Piscataway, NJ). The mononuclear cell fraction from the gradient was washed twice in CMF-HBSS and passed through a 35 mm sieve. The cells were suspended in growth medium consisting of a-MEM (pH 7.2, which was modified to contain 8.3 mM $NaHCO_3$ (Gibco BRL, Grand Island, NY)), 10% heat-inactivated fetal bovine serum (FBS, Hyclone, Logan, UT) and 2% antibiotic/antimycotic solution (Gibco BRL, Grand Island, NY) and seeded onto tissue culture dishes at a density of 1 x 10^6 cells/cm^2. A typical marrow cell yield was between 1-2 x 10^9 cells/animal, which varied with the size of the animal. The cells were incubated at 37°C in a humid incubator with 5% CO_2. After 24-48 h, nonadherent cells were removed and seeded in either 24-well cluster dishes at a density of 7.5 x 10^5 cells/cm^2 in growth medium which did or did not contain 10^{-8} M 1,25-$(OH)_2D_3$ (Biomol, Plymouth Meeting, PA) and TNFRsol protein (obtained as in Example 6). Cells were cultured for up to 10 days with medium changes every 48-72 h with growth medium that did or did not contain 1,25-$(OH)_2D_3$ and TNFRsol protein. Following 5 days of culture, the cells were fixed with formalin (3.7% for 10 min) and then stained for tartrate-resistant acid phosphatase (TRAP) as in Example 6. The number of osteoclasts (TRAP-positive cells containing 3 or more nuclei) was quantitated. Results are reported in Table 2.

Table 2

TNFRsol (ng/ml)	Osteoclasts/well[a]
0.00	214.83 + 14.22
0.01	68.83 + 6.28*
0.10	176.17 + 23.01
1.00	228.50 + 17.26
10.00	228.50 + 29.29
100.00	382.33 + 26.59*

a Each value represents the mean and standard error of 6 wells.
*$p<0.05$ compared to control group

EXAMPLE 8

Construction of TNFRsol-Flag Expression Vector

To facilitate confirmation of TNFRsol expression (without the use of antibodies), a bicistronic expression vector (pIG1-TNFRsolF) was constructed by insertion of an "internal ribosome entry site"/enhanced green fluorescent protein (IRES/eGFP) PCR fragment into the mammalian expression vector pGTD (Gerlitz, B. et al., 1993, Biochemical Journal 295:131). This new vector, designated pIG1, contains the following sequence landmarks: the E1a-responsive GBMT promoter (D. T. Berg et al., 1993 BioTechniques 14:972; D.T. Berg et al., 1992 Nucleic Acids Research 20:5485); a unique *BclI* cDNA cloning site; the IRES sequence from encephalomyocarditis virus (EMCV); the eGFP (Clontech) coding sequence (Cormack, et al., 1996 Gene

173:33); the SV40 small "t" antigen splice site/poly-adenylation sequences; the SV40 early promoter and origin of replication; the murine dihydrofolate reductase (dhfr) coding sequence; and the pBR322 ampicillin resistance marker/origin of replication.

Based upon the human TNFRsol sequence, the following primers were synthesized: 5'- TAGGGCTGATCAAGGATGG GCTTCTGGACTTGGGCGGCCCCTCCGCAGGCGGACCGGGG-3' (SEQ ID NO:3); and 5'- AGGGGGGCGGCCGCTGATCATCACTTGTCGTCGTCGTCCTTGTAGTCGTGCA CAGGGAGGAAGCGC - 3' (SEQ ID NO:4). The reverse primer contained the Flag epitope sequence (nucleotides 24-47 of SEQ ID NO:4) (Micele, R.M. et al., 1994 J. Immunol. Methods 167:279). These primers were then used to PCR amplify the TNFRsol cDNA. The resultant 1.3 Kb PCR product was then digested with *BclI* (restriction sites incorporated into primers, underlined above) and ligated into the unique *BclI* site of pIG1 to generate the plasmid pIG1-TNFRsolF. The human TNFRsol cDNA orientation and nucleotide sequence were confirmed by restriction digest and double stranded sequencing of the insert.

EXAMPLE 9

Construction of TNFRsol-non-Flag Expression Vector

In order to generate a non-Flagged expression vector (pIG1-TNFRsol), the 24-base DNA sequence encoding the eight amino acid FLAG epitope was deleted from the pIG1-TNFRsolF construct using the Quik Change mutagenesis kit (Stratagene). A 35-base primer, and its complement, with identity to the 19-base sequences flanking the FLAG sequence was synthesized and used to prime PCR using the plasmid as

template. The PCR product was digested with *DpnI* restriction endonuclease to eliminate the parental DNA, and the digested product was transformed into Epicurean XLI-blue E.coli cells. Sixteen ampicillin-resistant transformants were picked and the plasmid DNA was analyzed by restriction digestion. Ten of the 16 gave results compatible with deletion of the 24-base sequence. Precise deletion of the 24-base sequence was confirmed by DNA sequencing of pIG1-TNFRsol.

EXAMPLE 10

Isolation of a high-producing TNFRsol clone from AV12 RGT18 transfectants

The recombinant plasmid carrying the TNFRsol gene encodes resistance to methotrexate. In addition, the construct contains a gene encoding a fluorescent protein, GFP, on the same transcript and immediately 3' to the TNFRsol gene. Since high level expression of GFP would require a high level of expression of the TNFRsol-GFP mRNA, highly fluorescent clones would have a greater probability of producing high levels of TNFRsol. pIG1-TNFRsol and pIG1-TNFRsolF were used to transfect AV12 RGT18 cells. Cells resistant to 250 nM methotrexate were selected and pooled. The pool of resistant clones was subjected to fluorescence assisted cell sorting (FACS), and cells having fluorescence values in the top 5% of the population were sorted into a pool and as single cells. The high fluorescence pools were subjected to three successive sorting cycles. Pools and individual clones from the second and third cycles were analyzed for TNFRsol production by SDS-PAGE. Pools or

clones expressing TNFRsol at the highest level judged from Coomassie staining were used for scale-up and TNFRsol purification.

EXAMPLE 11

Large Scale TNFRsol Protein Purification

Large scale production of TNFRsol was done by first growing the stable clones in several 10 liter spinners. After reaching confluency, cells were further incubated for 2-3 more days to secret maximum amount of TNFRsol into media. Media containing TNFRsol was adjusted to 0.1% CHAPS and concentrated in an Amicon ProFlux M12 tangential filtration system to 350 ml. The concentrated media was centrifuged at 19,000 rpm (43,000 x g) for 15 minutes and passed over a SP-5PW TSK-GEL column (21.5 mm x 15 cm; TosoHaas) at a flow rate of 8 ml/min. The column was washed with buffer A(20 mM MOPS, 0.1% CHAPS, pH 6.5) until the absorbency (280 nm) returned to baseline and the bound proteins were eluted with a linear gradient from 0.1 M-0.3 M NaCl(in buffer A) developed over 85 min. Fractions containing TNFRsol were pooled and passed over a (7.5 mm x 7.5 cm) Heparin-5PW TSK-GEL column equilibrated in buffer B (50 mM Tris, 0.1% CHAPS, 0.3 M NaCl, pH 7.0). The bound protein was eluted with a linear gradient from 0.3 M-1.0 M NaCl (in buffer B) developed over 60 min. Fractions containing TNFRsol were pooled and passed over a 1 cm x 15 cm Vydac C4 column equilibrated with 0.1% TFA/H_2O. The bound TNFRsol was eluted with a linear gradient from 0-100% CH_3CN/0.1% TFA. Fractions containing TNFRsol were analyzed by SDS-PAGE and found to be greater than 95% pure and were

dialyzed against 8 mM $NaPO_4$, 0.5 M NaCl, 10% glycerol, pH 7.4. The N-terminal sequence of TNFRsol was confirmed on the purified protein. Mass spectral analysis and Endogylcosidase-F digestion indicates that TNFRsol is glycosylated.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Dou, Shenshen
Song, Ho Yeong

(ii) TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR FAMILY PROTEIN AND RELATED NUCLEIC ACID COMPOUNDS

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Eli Lilly and Company
(B) STREET: Lilly Corporate Center
(C) CITY: Indianapolis
(D) STATE: Indiana
(E) COUNTRY: USA
(F) ZIP: 46202

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Cantrell, Paul R.
(B) REGISTRATION NUMBER: 36,470
(C) REFERENCE/DOCKET NUMBER: P-11586A

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (317) 276-3885
(B) TELEFAX: (317) 276-5172

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 903 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..900

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATG AGG GCG CTG GAG GGG CCA GGC CTG TCG CTG CTG TGC CTG GTG TTG        48
Met Arg Ala Leu Glu Gly Pro Gly Leu Ser Leu Leu Cys Leu Val Leu
  1               5                  10                  15

GCG CTG CCT GCC CTG CTG CCG GTG CCG GCT GTA CGC GGA GTG GCA GAA        96
Ala Leu Pro Ala Leu Leu Pro Val Pro Ala Val Arg Gly Val Ala Glu
            20                  25                  30

ACA CCC ACC TAC CCC TGG CGG GAC GCA GAG ACA GGG GAG CGG CTG GTG       144
Thr Pro Thr Tyr Pro Trp Arg Asp Ala Glu Thr Gly Glu Arg Leu Val
        35                  40                  45

TGC GCC CAG TGC CCC CCA GGC ACC TTT GTG CAG CGG CCG TGC CGC CGA       192
Cys Ala Gln Cys Pro Pro Gly Thr Phe Val Gln Arg Pro Cys Arg Arg
    50                  55                  60

GAC AGC CCC ACG ACG TGT GGC CCG TGT CCA CCG CGC CAC TAC ACG CAG       240
Asp Ser Pro Thr Thr Cys Gly Pro Cys Pro Pro Arg His Tyr Thr Gln
65                  70                  75                  80

TTC TGG AAC TAC CTG GAG CGC TGC CGC TAC TGC AAC GTC CTC TGC GGG       288
Phe Trp Asn Tyr Leu Glu Arg Cys Arg Tyr Cys Asn Val Leu Cys Gly
                85                  90                  95

GAG CGT GAG GAG GAG GCA CGG GCT TGC CAC GCC ACC CAC AAC CGT GCC       336
Glu Arg Glu Glu Glu Ala Arg Ala Cys His Ala Thr His Asn Arg Ala
            100                 105                 110

TGC CGC TGC CGC ACC GGC TTC TTC GCG CAC GCT GGT TTC TGC TTG GAG       384
Cys Arg Cys Arg Thr Gly Phe Phe Ala His Ala Gly Phe Cys Leu Glu
        115                 120                 125

CAC GCA TCG TGT CCA CCT GGT GCC GGC GTG ATT GCC CCG GGC ACC CCC       432
His Ala Ser Cys Pro Pro Gly Ala Gly Val Ile Ala Pro Gly Thr Pro
    130                 135                 140

AGC CAG AAC ACG CAG TGC CAG CCG TGC CCC CCA GGC ACC TTC TCA GCC       480
Ser Gln Asn Thr Gln Cys Gln Pro Cys Pro Pro Gly Thr Phe Ser Ala
145                 150                 155                 160
```

```
AGC AGC TCC AGC TCA GAG CAG TGC CAG CCC CAC CGC AAC TGC ACG GCC        528
Ser Ser Ser Ser Ser Glu Gln Cys Gln Pro His Arg Asn Cys Thr Ala
                165                 170                 175

CTG GGC CTG GCC CTC AAT GTG CCA GGC TCT TCC TCC CAT GAC ACC CTG        576
Leu Gly Leu Ala Leu Asn Val Pro Gly Ser Ser Ser His Asp Thr Leu
            180                 185                 190

TGC ACC AGC TGC ACT GGC TTC CCC CTC AGC ACC AGG GTA CCA GGA GCT        624
Cys Thr Ser Cys Thr Gly Phe Pro Leu Ser Thr Arg Val Pro Gly Ala
        195                 200                 205

GAG GAG TGT GAG CGT GCC GTC ATC GAC TTT GTG GCT TTC CAG GAC ATC        672
Glu Glu Cys Glu Arg Ala Val Ile Asp Phe Val Ala Phe Gln Asp Ile
    210                 215                 220

TCC ATC AAG AGG CTG CAG CGG CTG CTG CAG GCC CTC GAG GCC CCG GAG        720
Ser Ile Lys Arg Leu Gln Arg Leu Leu Gln Ala Leu Glu Ala Pro Glu
225                 230                 235                 240

GGC TGG GGT CCG ACA CCA AGG GCG GGC CGC GCG GCC TTG CAG CTG AAG        768
Gly Trp Gly Pro Thr Pro Arg Ala Gly Arg Ala Ala Leu Gln Leu Lys
                245                 250                 255

CTG CGT CGG CGG CTC ACG GAG CTC CTG GGG GCG CAG GAC GGG GCG CTG        816
Leu Arg Arg Arg Leu Thr Glu Leu Leu Gly Ala Gln Asp Gly Ala Leu
            260                 265                 270

CTG GTG CGG CTG CTG CAG GCG CTG CGC GTG GCC AGG ATG CCC GGG CTG        864
Leu Val Arg Leu Leu Gln Ala Leu Arg Val Ala Arg Met Pro Gly Leu
        275                 280                 285

GAG CGG AGC GTC CGT GAG CGC TTC CTC CCT GTG CAC                        903
Glu Arg Ser Val Arg Glu Arg Phe Leu Pro Val His
    290                 295                 300
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 300 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Arg Ala Leu Glu Gly Pro Gly Leu Ser Leu Leu Cys Leu Val Leu
 1               5                  10                  15

Ala Leu Pro Ala Leu Leu Pro Val Pro Ala Val Arg Gly Val Ala Glu
            20                  25                  30
```

Thr Pro Thr Tyr Pro Trp Arg Asp Ala Glu Thr Gly Glu Arg Leu Val
35 40 45

Cys Ala Gln Cys Pro Pro Gly Thr Phe Val Gln Arg Pro Cys Arg Arg
50 55 60

Asp Ser Pro Thr Thr Cys Gly Pro Cys Pro Pro Arg His Tyr Thr Gln
65 70 75 80

Phe Trp Asn Tyr Leu Glu Arg Cys Arg Tyr Cys Asn Val Leu Cys Gly
85 90 95

Glu Arg Glu Glu Glu Ala Arg Ala Cys His Ala Thr His Asn Arg Ala
100 105 110

Cys Arg Cys Arg Thr Gly Phe Phe Ala His Ala Gly Phe Cys Leu Glu
115 120 125

His Ala Ser Cys Pro Pro Gly Ala Gly Val Ile Ala Pro Gly Thr Pro
130 135 140

Ser Gln Asn Thr Gln Cys Gln Pro Cys Pro Pro Gly Thr Phe Ser Ala
145 150 155 160

Ser Ser Ser Ser Ser Glu Gln Cys Gln Pro His Arg Asn Cys Thr Ala
165 170 175

Leu Gly Leu Ala Leu Asn Val Pro Gly Ser Ser Ser His Asp Thr Leu
180 185 190

Cys Thr Ser Cys Thr Gly Phe Pro Leu Ser Thr Arg Val Pro Gly Ala
195 200 205

Glu Glu Cys Glu Arg Ala Val Ile Asp Phe Val Ala Phe Gln Asp Ile
210 215 220

Ser Ile Lys Arg Leu Gln Arg Leu Leu Gln Ala Leu Glu Ala Pro Glu
225 230 235 240

Gly Trp Gly Pro Thr Pro Arg Ala Gly Arg Ala Ala Leu Gln Leu Lys
245 250 255

Leu Arg Arg Arg Leu Thr Glu Leu Leu Gly Ala Gln Asp Gly Ala Leu
260 265 270

Leu Val Arg Leu Leu Gln Ala Leu Arg Val Ala Arg Met Pro Gly Leu
275 280 285

Glu Arg Ser Val Arg Glu Arg Phe Leu Pro Val His
290 295 300

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TAGGGCTGAT CAAGGATGGG CTTCTGGACT TGGGCGGCCC CTCCGCAGGC GGACCGGGG 59

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGGGGGGCGG CCGCTGATCA TCACTTGTCG TCGTCGTCCT TGTAGTCGTG CACAGGGAGG 60

AAGCGC 66

WE CLAIM:

1. A substantially pure protein comprising an amino acid sequence which is amino acids 34-195 of SEQ ID NO:2.

2. The substantially pure protein of claim 1 comprising an amino acid sequence which is amino acids 30-300 of SEQ ID NO:2.

3. The substantially pure protein of claim 1 comprising an amino acid sequence which is SEQ ID NO:2.

4. An isolated nucleic acid compound encoding the protein of Claim 1, or a sequence complementary to said compound.

5. An isolated nucleic acid compound encoding the protein of claim 2, or a sequence complementary to said compound.

6. An isolated nucleic acid compound encoding the protein of Claim 3, or a sequence complementary said molecule.

7. An isolated nucleic acid compound encoding a protein having osteoclast differentiation-inhibiting activity, wherein said compound hybridizes to a nucleic acid compound as set forth in SEQ ID NO:1 under high stringency conditions.

8. The isolated nucleic acid compound of claim 7 that is at least 75% identical to a nucleic acid compound as set forth in SEQ ID NO:1.

9. A vector comprising the isolated nucleic acid compound of Claim 4.

10. The vector of Claim 9, wherein said isolated nucleic acid compound is SEQ ID NO:1 operably-linked to a promoter sequence.

11. A host cell transformed with a vector of Claim 9.

12. A host cell transformed with a vector of Claim 10.

13. A method for constructing a recombinant host cell having the potential to express a protein comprising an amino acid sequence consisting of amino acids 34-195 of SEQ ID NO:2, said method comprising introducing into said host cell by any suitable means a vector of claim 10.

14. A method for expressing a protein identified herein as SEQ ID NO:2 in a recombinant host cell of Claim 13, said method comprising culturing said recombinant host cell under conditions suitable for gene expression.

15. A method for identifying compounds that bind a protein identified herein as SEQ ID NO:2, comprising the steps of:

a) admixing a substantially purified preparation of a protein comprising SEQ ID NO:2 with a test compound; and

b) monitoring by any suitable means a binding interaction between said protein and said compound.

16. A method, as in Claim 15 wherein said protein is identified herein as SEQ ID NO:2.

17. An antibody that selectively binds to a protein identified herein as SEQ ID NO:2, or fragment thereof.

18. The isolated nucleic acid of Claim 4 as set forth in SEQ ID NO:1.

19. The isolated nucleic acid of Claim 4 corresponding to nucleotides 88-900 of SEQ ID NO:1.

20. The isolated nucleic acid of Claim 4 corresponding to nucleotides 102-585 of SEQ ID NO:1.

ABSTRACT

The invention provides isolated nucleic acid compounds, proteins, and fragments thereof, said proteins being related to the family of tumor necrosis factor receptors. Also provided are vectors and transformed heterologous host cells for expressing the protein and a method for identifying compounds that bind and/or modulate the activity of said proteins.